MODELING STUDY OF PCBs IN THE HOUSATONIC RIVER PEER REVIEW

Modeling Framework Design
Final Written Comments

Peter Shanahan
HydroAnalysis, Inc.
May 28, 2001

RESPONSE TO CHARGE FOR THE HYDRODYNAMIC MODELING PEER REVIEW

I. General Overview of Response

This document provides my review of the proposed modeling study of contamination of the Housatonic River by polychlorinated biphenyls (PCBs). It is based on a review of the Modeling Framework Design (MFD) (Beach *et al*., 2000a), its accompanying Quality Assurance Project Plan (QAPP) (Beach *et al*., 2000b), and various supporting documentation and data provided by the U.S. Environmental Protection Agency (EPA). My views were further influenced by the Peer Review Workshop held April 25 and 26, 2001 in Lenox, Massachusetts.

CONCEPTUAL MODEL

A major concern discussed at length during the Peer Review Workshop is the soundness of the conceptual model of the Housatonic River system developed by EPA. Chapter 3 of the MFD does not provide a succinct coherent summary of the conceptual model of the Housatonic River system. Rather, it provides a lengthy discussion of all phenomena that may be influential and the various data available to describe those phenomena. In this sense, it is more of a "data dump" than a conceptual model. There is little provided in the way of data analysis, calculations, simple "back-of-the-envelope" models, and dimensionless scaling analysis based on the data. What is particularly missing is a sense of priority: that is, of which processes are of paramount importance and must be captured in the model and which are secondary and can be omitted without compromising predictive ability.

The weakness of the conceptual model comes to the fore in making decisions regarding how to model the hydrodynamics and sediment transport of the floodplain. This is the subject that garnered the most attention from the Peer Review Panel during their April Workshop. It is important because of the significant computational difficulties it poses. The hydrodynamic interaction of the river channel with the distal floodplain is proposed by EPA to be handled in an approximate fashion using a dual-grid approach that does not conserve momentum. However, a fully rigorous two-dimensional model of the river channel and floodplain would carry a punishing and probably infeasible computational burden at least with the EFDC model. As it is, the approximate approach put forward by EPA is still so computationally intensive as to preclude systematic uncertainty analysis.

Lack of a clear conceptual model for the floodplain prevents making clear recommendations for an alternative hydrodynamic model. The conceptual model put forward in the MFD and supplementary responses to the peer review panel by EPA seems to be the following: Most PCB-contaminated sediments are found in the river channel and Woods Pond (accounting for 37% and 17%, respectively, of the mass of PCBs in the Primary Study Area (PSA)). An additional 17% of the PCB mass lies in the proximal floodplain. These fractions of the PCB mass, which amount to about 71% of the total in the PSA, lie in areas within or directly adjacent to the river channel that can modeling straightforwardly. The remaining 29% of the PCB mass lie in the distal floodplain. The distal floodplain is inundated only infrequently, by storms occurring less often than about once in every 5 years. When inundated, river flow over the floodplain is slowed by vegetation, with the result that sediment is deposited on the floodplain. However, as stated on page 3-49 of MFD, "Erosion of the floodplain can occur under extreme high flow events and from bank slumping."

The specific basis for this last statement regarding the effect of extreme high flow events is not stated in the MFD and is not apparent from the Workshop discussions and other materials provided by EPA. Sedflume erosion studies completed by Gailani *et al.* (2000) did not include floodplain samples, so the erosional properties of floodplain soils are not known in detail. The storm events monitored in detail during 1999 do not seem to include an event large enough to have inundated the distal floodplain, much less to have eroded it, so there are no observations of distal floodplain erosion. While it seems logical that erosion of floodplain soils might occur during large floods, the heavy vegetative cover would likely minimize the extent and degree of such erosion. Finally, even if some erosion were to occur, it is not a foregone conclusion that the mass of PCBs mobilized by such an event would be significant. After all, the entire distal floodplain area contains only 29% of the PCB mass, PCB concentrations in floodplain soils are less than in other areas according to Table 3-8 of the MFD, and the floodplains are a net depositional area. In the final analysis, there is not a clear conceptual model of the importance of the distal floodplain to the eventual recovery of the Housatonic River.

The import of understanding the distal floodplain as a source is great. Were the floodplain a wholly or predominantly one-way sink for PCBs, modeling could be greatly simplified. There would be no need to model the distal floodplain per se. Rather, it could be simply considered as a sink for PCBs from the river system. Remediation alternatives for floodplain deposits could be evaluated as a separate issue, based on risk assessment rather than modeling. The Housatonic River could then be considered as an essentially one-dimensional channel, without all of the complexities and computational burden engendered by the distal floodplain. A simpler hydrodynamic model, possibly a one-dimensional model, would suffice, enabling greater computational flexibility. With fewer computational demands, the modeling study could address other concerns raised by the peer review panel, specifically, the completion of uncertainty analyses that could place the model results in perspective, and the simulation of multiple fractions of PCBs rather than simply total PCBs.

It is my recommendation that the role of the floodplain be better defined before the EPA commits to an overly complicated and computationally burdensome hydrodynamic model. Sedflume studies should be conducted on floodplain soils to understand the propensity of those soils to erode. Simple calculations should be completed to estimate stream velocities and shear stresses in the floodplain during large storm events. With that information, estimates of the mass of PCBs released from the distal floodplain during large storm events should be developed and compared with mass estimates for existing channel deposits and stormflow suspended solids. That comparison will provide an indication of the significance of floodplain-derived PCBs and enable decisions as to the need for and approach to modeling the distal floodplain. Only then should the hydrodynamic modeling approach be committed. At present, in the absence of a coherent model of the role and importance of the distal floodplain, decisions as to hydrodynamic modeling are premature.

MODEL COMPLEXITY

Despite lip service to model parsimony in Section 3.1.1 of the MFD, the models proposed for this study appear to have been developed with the philosophy that the more detailed and complex the formulation, the better. I see intrinsic difficulties in this approach, although my aversion is more visceral than analytical. Fortunately, Bruce Beck has provided an analysis of the issue that is both philosophical and rigorous in two papers (Beck, 1987; Beck and Halfon, 1991). The problems he identifies can be generally described as an inability to develop a proper model structure, uncertainty in the model parameters, and poor predictive capability.

The pertinence of the issue of proper model structure is demonstrated in a recent review of the models developed to assess PCB contamination in the Fox River in Wisconsin (Tracy and Keane, 2000). A physically-based (as opposed to empirical) model of the Fox River was developed by the Wisconsin Department of Natural Resources (WDNR), calibrated to field data for PCBs and suspended solids, and apparently used as the basis for a draft Superfund feasibility study (FS) of alternative strategies to remediate the river. Alternative cleanup strategies range from no action to dredging and contained disposal of contaminated sediments at an estimated cost of $720 million (WDNR, 1999).

Subsequent to the publication of the models and draft FS, the WDNR model has been examined in detail by consultants to the industries identified as Potentially Responsible Parties (PRPs). The consultants identified a serious structural flaw in the model. Numerical dispersion arising from the representation of PCB transport in the sediments produces a large artificial flux of PCBs from deep sediment layers to the sediment surface in the WDNR model. As the result of this and other differences in the models, the WDNR model predicts that 70% more PCBs will be discharged from the Fox River than does an alternative model developed by the PRPs' consultants. WDNR has corrected its model, but has not yet reported the effect on the model predictions. There appears to be no consensus among the various parties on a best model or best modeling approach and, as indicated by Dr. Lick during the Peer Review Workshop, estimates of bed erosion between the parties differ by two orders of magnitude.

The specific flaw identified in the WDNR Fox River model is known to the modelers working on the Housatonic River and will not be repeated. Moreover, the Housatonic River modelers also know the recommendations of the Fox River Peer Review Panel, although the study is cited only in the EPA Response to Peer Review Panelist Questions and not in the MFD or QAPP. Regardless of the specific flaw in the WDNR model and its correction, the Fox River example reveals a fundamental failing in the application of these highly complex models: the models are over-parameterized such that even an intrinsically flawed model can be "calibrated" to field data. I suspect that with the number of parameters included in the AQUATOX model, it could be satisfactorily calibrated to any time-series data set, including the Dow Jones Industrial Average.

Beck's review of model uncertainty leaves me pessimistic—he states for example that

"Overparameterization seems both intrinsic and an intractable problem" (Beck and Halfon, 1991). He also makes clear that a "physics-based mechanistic" approach is hardly a panacea (Beck, 1987). It is noteworthy that he cites the predecessor model of AQUATOX as an example of the misguided physics-based approach. In essence, Beck argues that these models cannot be tested by the scientific method because they cannot be shown to be false. There are too few field data to disprove the many subordinate hypotheses and parameters embedded in these complex models. Thus, one cannot test hypotheses as is incumbent with the scientific method.

Despite this overall pessimism, Beck (1987) does provide some optimistic observations. In particular, he distinguishes simulation of hydrology (and by implication, hydrodynamics) from simulation of water quality. He cites the longer and more intensive study of hydrologic processes compared to the more ephemeral attention to water quality (which has shifted attention over the years from BOD/DO, to eutrophication, to acid precipitation, and now to toxic chemicals). As a result of more intensive historical examination, the hydrological and hydrodynamic processes are better understood, more certainly parameterized, and better identified. In the context of the Housatonic River modeling framework, the HSPF and hydrodynamic portion of the EFDC fall within this class of more certain models. The comparatively empirical AQUATOX and sediment transport portion of EFDC are within the less certain category.

Another seemingly pessimistic aspect of Beck's analysis is the fact that his focus is primarily on eutrophication modeling, which is only a subset of the modeling exercise proposed here and thus less complex. However, as Beck (1987, Figure 15) illustrates, the eutrophication problem has some intrinsic difficulties that may not be shared by the PCB problem. Specifically, the prediction of eutrophication involves the translation of relatively steady meteorological and nutrient loading forcing functions into episodic algal blooms—an abrupt transient response that bears little resemblance to the character of the forcing functions. The PCB problem can be idealized as a relatively better behaved problem: namely the exponential depletion and burial of mass over decadal time scales. Indeed, with this conceptual model, I wonder if a calibrated exponential decay coefficient for PCB loss could be as reliable a predictor as the modeling effort proposed here. A flaw in the exponential conceptual model is, of course, the potentially great influence of unusually high flow events. However, the predictive ability of the proposed modeling framework for high flow events is perhaps the single greatest uncertainty in the model.

In the final analysis, AQUATOX appears to be overly complex for the task it was chosen to do: modeling the bioaccumulation of PCBs within the food chain. As explained by Mr. Endicott during the Peer Review Workshop, AQUATOX simulates the entire aquatic ecosystem. This incorporates and adds to the already complex eutrophication model about which Beck is so pessimistic in his analysis of model uncertainty. The alternative, as also explained by Mr. Endicott, is the simpler food-chain models that have already been used in other riverine PCB models. This alternative eliminates much of the duplication and potential conflict of modeling PCBs by two different approaches with EFDC and AQUATOX. The MFD, QAPP, and EPA Responses provide no evaluation of why the AQUATOX approach is superior to this more established alternative. One can in fact view this as a failure of the Quality Assurance process: the QAPP accepts the models chosen in the MFD as a given, but should in fact provide for quality assurance evaluation of the process by which particular modeling codes are chosen.

Overall, the considerable complexity, and possible "over-complexity," of the proposed models creates challenges for model calibration and necessitates an explicit evaluation of the confidence that can be placed in the model predictions. These topics are further discussed in the following.

MODEL CALIBRATION

Beck's analysis makes clear that a "good" calibration in and of itself does not guarantee that the model has been defined correctly or that its parameter values are reasonable. Moreover, simple statistical measures may be misleading as was demonstrated by the apparently successful calibration of the incorrect Fox River model.

A partial remedy to the calibration issue is achieved in the Housatonic River MFD and QAPP by disaggregating certain process and calibrating those separately. For example, completion of an independent sediment erosion study (Gailani *et al*., 2000) provides some assurance that even if the entire modeling framework cannot be calibrated and verified, important sediment processes can be adequately represented. The more such "subcalibrations" can be completed, the more confidence can be ascribed to the overall model. This approach seems to be endorsed by the Fox River Peer Review Panel as well (Tracy and Keane, 2000). Thus, I would encourage close scrutiny of the models for subprocesses that can be isolated and tested. The evaluation of sediment erosion using Sedflume is a critical such analysis, but should be extended to distal floodplain soils as well as the in-channel samples analyzed by Gailani *et al.* (2000).

The following discusses the calibration for the individual model components: HSPF, EFDC, and AQUATOX.

HSPF. Given its long history of use, and the fact that it has been tested and found successful in at least one post-audit (Hartigan, 1983), I have more confidence in the use of HSPF than of the other models. Nonetheless, I found the description of the HSPF calibration in Section 4.5 of the QAPP to be inadequate. The text is a litany of the parameters that will be adjusted, and the sequence in which they will be adjusted, but provides little information on the field data to which the data will be calibrated. It is obvious however that the available data will allow only gross calibration of the model based primarily on mainstem hydrologic stations. Data collected in 1999 are apparently the only tributary data available, although the MFD's description of the datasets is sparse. The Supplemental Investigation Workplan (Weston, 2000, Figure 5.3-1) shows a limited number of tributary stations. This data set will be used for calibration, apparently leaving no tributary dataset available for validation. Given the history of use of HSPF this seems acceptable although undesirable.

EFDC. The application of EFDC (or other hydrodynamic and sediment transport codes) to the Housatonic River is clearly a difficult undertaking. The hydrodynamic regime includes transitions between at least three general modes of flow:

- in-bank flow within the meandering river during most of the time;
- out-of-bank flow in the also meandering proximal flood plain during flood events of less-than-approximately-5-year recurrence intervals; and
- widespread flow within the relatively straight floodplain during large flood events (greater-

than-approximately-5-year recurrence).

Further complexity affects the two flood-flow modes due to the extensive floodplain vegetation, which reduces flow velocities and enhances sediment deposition. As discussed above, data are insufficient to describe and understand the distal floodplain, and the conceptual model of the floodplain is incomplete.

Presuming the plan to use the EFDC model for hydrodynamics and sediment transport goes forward, the dataset for calibration of the EFDC model is insufficient to test a critical aspect of the model: its ability to accurately predict the effect of high-flow events. The potential impact of future high flow events is a matter of speculation at other riverine PCB contamination sites, but has nonetheless been held out as a rationale for expensive dredging remediation alternatives. Until a sufficiently extreme flow event presents itself and is thoroughly monitored, the predictive abilities of the Housatonic River model will remain in doubt. Completing Sedflume studies on the floodplain soils can reduce this uncertainty, but the EPA should retain the ability to thoroughly monitor a large storm event should one present itself.

AQUATOX. The AQUATOX model is a classic example of an overparameterized model. The QAPP lists numerous statistical and qualitative measures that will be used to assess the calibration of the AQUATOX model, but these will provide no assurance that the model calibration is unique or correct. There are simply too many parameters and too few data to calibrate the model in a completely satisfactory way. I see no escape from this dilemma other than that the model predictions must be qualified by explicit calculation of the model uncertainty so that decision makers can judge the confidence with which the model results can be used.

UNCERTAIN PREDICTIVE ABILITY

Prognostic simulations included within the model objectives include estimates of the effects of selected remedial actions, of natural recovery, and of extreme storm events. Remedial alternatives in the case of river sediments necessarily include dredging. Since opinions on dredging run the gamut from a remedial panacea that is practically pre-ordained to counterproductive folly that only spreads and worsens the contamination, it is important to foresee as a part of the modeling framework design how dredging will be evaluated with the model. The representation of dredging, though a critical factor in the eventual model-based decision-making, is absent from the MFD. I am pessimistic that a fair assessment of the effects of dredging will be accomplished without its prior explicit specification. How will sediment disruption and collateral release during dredging be modeled? Is this modeling framework adequate to address that question? I fear that if these questions are not answered upfront, the eventual model application will founder on disagreements over its application.

There is also the considerable uncertainty, alluded to earlier, in the use of the models to forecast behavior of the system at two extremes: the high-flow event, and the system response over decades. The short record available for calibration, and the absence of high-flow events from that record, are ominous for both of these capabilities. By definition, it is unlikely that a high-flow event will be observed during the data-collection period, and it is impractical to extend the observation period to decades. Thus, neither of these critical needs is likely to be addressed. This is not just an issue for model calibration but also of model identification. Specifically, there is considerable uncertainty that the EFDC model can adequately represent the transition from in-bank flow within the meandering stream channel to overbank flow through the floodplain.

How then, in the absence of adequate data, can the uncertainty in these extreme predictions be factored into decision-making? It seems that there is only one possibility, namely to include within the modeling exercise explicit tasks to estimate and report the uncertainty bounds on the model estimates. This task is given only token attention in the MFD, and even then indicated as an optional task that "may" be done. In the absence of explicit evaluation of uncertainty, the inevitable result is that described by Beck and Halfon (1991):

> *There must have been many occasions on which a large model, confronted with hopelessly inadequate data, has been accepted as an appropriate tool for prediction, and without any attempt at quantification of the error attaching to the predictions so obtained.*

In effect, without explicit evaluation of uncertainty, the model is implicitly portrayed as certain. For example, the tracing of a single solid predicted-results line on a graph, without error bounds, would represent an implicitly certain forecast. The risk is that unnecessary exposure to PCBs and unwarranted expenditures would result from the selection of a predicted outcome that does not in actual fact differ from an unselected alternative.

The MFD indicates that the complexity of the models, and particularly the EFDC code, precludes rigorous evaluation of uncertainty. In its place, sensitivity tests are proposed. I found the description of the approach to be inadequate however. It seems to me that it is at least as important to communicate to decision makers the level of uncertainty in the model predictions as it is to communicate the predictions. Thus, I believe more explicit and formalized procedures for the calculation and communication of uncertainty should be incorporated into the MFD.

Some of the other members of this panel have asked about a contingency plan in the event the models are found inadequate to meet the study objectives. While I do not share concerns about this issue per se, I have a closely related question, namely: With what confidence can the models be used to answer the questions posed by the study objectives? In essence, I assume the modelers can get an answer—my concern is the utility and reliability of that answer. That reliability can only be known to the decision makers if the modelers explicitly calculate the uncertainty and reliability of their model predictions. It is essential to add this task to the modeling study in my opinion. This also implies that it is essential to find a simplified alternative to the EFDC model to make an uncertainty analysis computationally feasible.

II. Response to Peer Review Questions

In considering the foregoing general issues and evaluating the EPA documents, the Peer Review Panel shall give specific consideration to the following questions. As modeling activities proceed, additional specific questions may be identified the panel to address.

A. *Modeling Framework and Data Needs*

1. ***Do the modeling frameworks used by EPA include the significant processes affecting PCB fate, transport, and bioaccumulation in the Housatonic River; and are the descriptions of these processes in the modeling framework(s) sufficiently accurate to represent the hydrodynamics, sediment transport, PCB fate and transport, and PCB bioaccumulation in the Housatonic River?***

The presumption of this question is that "accurate" models and process descriptions are possible. With processes of the complexity and uncertainty as those at issue, a focus on achieving accuracy is misplaced. Rather, it should be recognized that the models are necessarily inaccurate. The focus, therefore, should be on quantifying the uncertainty of the model predictions through as rigorous an uncertainty analysis as possible. Information on the uncertainty of the model predictions then needs to be provided to decision makers and the public along with the model predictions so that truly informed decisions can be made.

The computational burden of the EFDC model is clearly an impediment to achieving the goal of quantifying uncertainty. As indicated by EPA during the Peer Review Panel Workshop, the EFDC model requires too much computer time to be subject to a quantitative uncertainty analysis. Unless ways can be developed to remedy this defect, the inability to assess uncertainty makes EFDC an unsuitable tool for this study. Nonetheless, there may be ways to remedy the defect including developing a simplified version of the EFDC model specifically for uncertainty analysis or by completing a rigorous uncertainty analysis on a representative subreach of the model. In any case, as indicated in my General Overview above, I believe that any decision to use the EFDC code is premature until more data are collected and a coherent conceptual model of the distal floodplain is developed.

Issues of uncertainty aside, the models need at least to attempt to include all significant processes with an appropriate degree of accuracy. I use the words "appropriate degree of accuracy" in conscious distinction from Question 1's phrase "sufficiently accurate." The term sufficiently accurate implies that insufficient accuracy is unacceptable, but excessive accuracy is irrelevant. In fact, greater "accuracy" in representing physical processes usually implies greater model complexity, more parameters, and a greater computational burden. These by-products of accuracy can be as detrimental to the overall success of the modeling program as insufficient accuracy, and are to be avoided as well.

With this in mind, I am concerned that all of the models may suffer from some degree of "over-accuracy." I am least concerned in this regard with HSPF. Although complicated, HSPF has a long history of use and an experience base in choosing parameter values. The intended use of the HSPF model for the Housatonic River study seems to be in flux, although in a favorable direction. The MFD and numerous responses to the Panel's question clearly indicate that HSPF was intended to be used to generate PCB loadings. We learned at the Peer Review Panel Workshop that this was not the plan after all, and that PCB loadings would be generated by coupling HSPF flow predictions with PCB field data measurements. The exact nature of this procedure has apparently not yet been documented in writing or otherwise defined in detail and therefore has not been subject to this peer review. Nonetheless, the decision to not use HSPF to generate PCB loads eliminates the most uncertain and speculative aspect of the MFD's plan for HSPF. Thus, overall, the use of HSPF satisfies the goal of "accuracy."

The EFDC model is a conundrum as far as process representation. The model includes a great many processes, but most are physically-based and reasonably well established. Thus, I am for the most part comfortable with the representation of individual processes and the relative absence of poorly defined coefficients. At a larger scale, however, the aggregation of so many processes in a fine-grid model has created an unfortunately large computational burden. As a result, in the net, the model is over-accurate to the point that the model's overall utility is diminished.

The exception to my general characterization of the EFDC code is the proposed dual-grid scheme for representing channel-distal floodplain interaction. The proposed approach is essentially experimental and there is little basis to judge whether the accuracy of this process in the model is sufficient. The fact that momentum is not conserved between the channel and distal floodplain represents a significant compromise in theoretical accuracy, but the impact on practical accuracy (i.e., predictive ability) is uncertain. In the net, I view the application of EFDC to this complex river system to be experimental and the accuracy for this application to be at best uncertain.

The AQUATOX model is highly overparameterized, as discussed in my general overview. In this sense, I view it as "over-accurate." The effect of this is to diminish the confidence in the model parameterization and calibration. That said, I was assured in the Peer Review Workshop that the model runs efficiently enough to be used in formal uncertainty analysis. Completion of such an analysis would mitigate my concerns about the accuracy of this model. Nonetheless, as stated in my General Overview, I believe simpler models are preferable to AQUATOX.

The modeling framework also incorporates algorithms for linking the various model components. None of the models are consistent in their representation of the pertinent state variables, and translation is required to transfer information from HSPF to EFDC and from HSPF and EFDC to AQUATOX. This seems an intrinsically inaccurate undertaking that adds uncertainty to the model results. To the extent possible, the conversion algorithms should be individually tested and validated, the their contribution to model uncertainty quantified. Better still would be the use of other models in which conversion of state variables in unnecessary.

2. Based upon the technical judgment of the Peer Review Panel:

a. ***Are the modeling approaches suitable for representing the relevant external force functions (e.g.. hydraulic flows, solids and PCB loads, initial sediment conditions, etc.), describing quantitative relationships among those functions, and developing quantitative relationships between those functions and PCB concentrations in environmental media (e.g., water column, sediments, fish and other biota, etc.)?***

Modeling approaches for forcing functions appear, for the most part, to be reasonable and appropriate at at least a conceptual level. At an operational level, some of the forcing functions for EFDC and AQUATOX are generated in HSPF and EFDC and must be converted as a part of the model linkage algorithms. The model linkages are not well defined in the MFD, but the presumption seems to be that empirical correlations can be used to define conversion formulas. As discussed under Question 1 above, this procedure is fraught with potential difficulties and needs to be avoided if possible, but thoroughly tested if unavoidable.

I am also concerned with respect to forcing functions is the representation of remediation. This affects both the upstream boundary condition and the simulation of future scenarios in the PSA. As far as the upstream boundary, we were told that boundary conditions might span the range of assuming upstream remediation to be 100% effective to 0% effective. Such uncertainties introduced at the upstream boundary could overshadow all of the model's predictions. Similar concerns affect the representation of remedial alternatives within the PSA, as discussed in my General Overview above. The Modeling Framework is incomplete until these very important model forcing functions are determined.

b. ***Are the models adequate for describing the interactions between the floodplains and the river?***

This particular question is pertinent to only the EFDC model: neither HSPF nor AQUATOX address the floodplains. The application of EFDC (or other hydrodynamic and sediment transport codes) to the Housatonic River is clearly a difficult undertaking that is complicated by different modes of flow within the channel, in the proximal floodplain, and in the distal floodplain. As discussed in my General Overview, there does not seem to be a coherent conceptual model of the floodplain to enable decisions on how it should be modeled.

The proposed application of the EFDC model to this complicated flow regime is essentially experimental. The proposed solution for modeling is the creation of a dual-grid model: a curvilinear grid along the river channel and proximal floodplain, and a separate linked-grid for the distal flood plain. The linkage between the two grids is not described in detail in the MFD, its appendices, or the EPA's Responses to Peer Review Panelist Questions. The EPA's Responses indicate that the dual-grid scheme will not conserve momentum.

The uncertainty of the dual-grid approach is implicit in the EPA's indication that the linkage mechanism will be tested, although the character of that testing also is not defined in detail. It appears that this testing will be a comparison against a single-grid model rather than field data, which are lacking. With respect to field data, there do not appear to have been hydraulic measurements made within the floodplain against which to test the model of floodplain flow. Sediment erosion tests by Gailani *et al.* (2000) were also restricted to the river channel. Thus, the floodplain linkage algorithm appears to be uncertain and there do not appear to be field data to test it.

As stated in my General Overview, I am not convinced there is a need to model the distal floodplain explicitly. Nonetheless, if one assumes that a model incorporating the distal floodplain is needed, it is incumbent to find an acceptable alternative to the proposed model. Unfortunately, although there seem to be problems with the dual-grid approach, there do not appear to be clearly superior alternatives. My sense is that a finite-element alternative, such as HSCTM2D, might present greater flexibility in structuring a grid that provides needed detail in the river channel and proximal floodplain, but lesser detail (and computational efficiency) for the distal floodplain. While the EPA's responses (pg. 14) indicate that HSCTM2D is considerably slower than EFDC, no details on the comparison are given and I wonder about the relevance of that comparison to the unusual Housatonic River configuration. I also am unable to shake my concern, which is possibly unfair, that alternatives to EFDC were not seriously considered, but that the model's availability and modeler's prior experience with the code preordained its selection.

In the final analysis, the answer to Question 2b is "I do not know." Changing this answer requires the type of conceptual model development described in my General Overview above.

c. ***Are the models adequate for describing the impacts of rare flood events?***

The question is relevant to HSPF and particularly EFDC. It has limited relevance to AQUATOX given the longer time step and limited spatial domain (no overbank areas) of that model. With respect to HSPF, I see no fundamental limitation in HSPF to capturing rare flood events.

With EFDC, the complexity of the Housatonic River and the use of the approximate linked-grid approach raise doubts as to the ability of the model to capture rare flood events in the same way they raise doubts as to describing channel-floodplain interaction. The lack of intensive data collection during a truly high-flow event (i.e., with flow within the distal floodplain) impedes the modeler's ability to test the model's predictions for high-flow events. Moreover, a lack of Sedflume data from the floodplain areas means that there are few data to specify sediment-related parameters for the distal floodplains during high-flow events.

As with Question 2b, the answer to this question is "I do not know" in the absence of more data and a coherent conceptual model for the distal floodplains as discussed in the General Overview.

d. ***Are the models adequate for discriminating between water-related and sediment-related sources of PCBs to fish and other biota?***

This question is relevant primarily to AQUATOX. Given the many transformation pathways included in AQUATOX, the model is fundamentally able to discriminate between water-related and sediment-related sources of PCBs to biota. However, the accuracy of the predictions is a function of the parameter values chosen for use in the model. As discussed in my General Overview, there are too few data and too many parameters in AQUATOX to ensure reliable determination of parameter values. Thus, while the model algorithms may be adequate, the model predictions can be expected to be highly uncertain. Unless this uncertainty is quantified and communicated to decision makers, I would not consider the model application to be adequate.

The sediment transport predictions by EFDC also relate to this question. Setting aside the reservations expressed above concerning floodplain and large-flood predictions, the EFDC model should be adequate for predicting sediment movement in the river channel. Again, these predictions are subject to parameter uncertainty. In this case, the parameter uncertainty is compounded by the uncertainties in modeling channel-floodplain interaction and large flood events. Here again, unless uncertainty is quantified and communicated, I would not consider the model application to be adequate.

3. Again, based upon the technical judgment of the Panel, are the spatial and temporal scales of the modeling approaches adequate to address the principal need for the model - producing sufficiently accurate predictions of the time to attain particular PCB concentrations in environmental media under various scenarios (including natural recovery and different potential active remedial options) to support remedial decision-making in the context described above in the Background section? If not, what levels of spatial and temporal resolutions are required to meet this need?

As with some of the other questions above, the issue may not be: Are the scales adequate? (i.e., fine enough spatially or short enough temporally), but, Are they too small? Short time steps and fine spatial detail increase the computational burden and reduce the ability to assess model uncertainty.

For HSPF, the spatial scale (hydrologic subbasins) and time scale (hours) are logical and commonplace in hydrologic analysis.

For the EFDC code, this particular question is premature inasmuch as the MFD and EPA responses give no explicit recommendation as to the grid size. The EPA's response to questions 52, 54, and 118 implies a channel grid size that is relatively small: $\Delta y \cong 25$ feet (three grids across the channel) and, very roughly, $\Delta x \cong 200$ feet. Small spatial elements create a double penalty in computational burden: they require the time step to be reduced for computational stability and also increase the number of computations needed to cover the spatial domain. Thus the selection of the EFDC spatial resolution is a critical decision. The longitudinal distance is not excessively small relative to field data density or the Cartesian grid size recommended by GE (20 meters = 66 feet). On the other hand, the lateral size appears to exceed the resolution of most of the field data and the EPA's responses indicate no intent to model gravel and point bar deposits within the channel. These considerations suggest that a single cross-sectional element may suffice, which would reduce the computational burden. A further improvement could be the alternative discussed at length by the peer review panel—a one-dimensional model. In either case, the modelers should investigate the sensitivity of the predictions to grid size and use as large a size as possible. Given that the eventual outcome of the model is prediction of decadal-scale recovery, loss of spatial detail in the hydrodynamic and sediment transport code should not compromise predictive ability so long as the overall flux of PCBs to and from the sediment is represented reasonably.

For the AQUATOX code, presuming a simpler alternative is not substituted, the spatial scale and temporal scales are appropriately large. This reduces the computational burden, allowing uncertainty analysis, but still captures appropriate system dynamics on decadal scales. My one concern with respect to the spatial scales of the AQUATOX model is that for large-scale box models of this type, the dispersion coefficients determined from the hydrodynamic model results should be adjusted to account for the implicit dispersion in the large-box AQUATOX elements (see Shanahan and Harleman, 1984). The discussion in the MFD implies that it is simple and straightforward to aggregate flows and fluxes from EFDC for use in AQUATOX. There are in fact considerable subtleties in such aggregation including the introduction of artificial or erroneous fluxes.

4. Is the level of theoretical rigor of the equations used to describe the various processes affecting PCB fate and transport, such as settling, resuspension, volatilization, biological activity, partitioning, etc., adequate, in your professional judgment, to address the principal need for the model (as defined above)? If not, what processes and what resolution are required?

The focus of this question is on theoretical rigor when many of the equations for hydrologic processes in HSPF, sediment resuspension, settling, and transport in EFDC, and biotic interaction in AQUATOX are empirical rather than theoretical. Indeed, one could argue that the only consistently theoretically rigorous aspect of the models is the hydrodynamic model in EFDC. Ironically, EPA plans to violate this uniquely rigorous computation in the dual-grid linkage between the stream channel and distal floodplain by failing to conserve momentum. EPA did not justify the necessity and/or benefit of this deviation from theory. Unless the computational savings are substantial, and the deviation can be demonstrated to be harmless, this approach should be avoided.

The EFDC code also deviates from theoretical rigor by lumping all PCBs into a single state variable and representing their highly disparate adsorptive and other properties with single, empirically determined values. Again, the computational burdens on EFDC appear to be influential in this question. It would seem far more defensible to divide total PCBs into coherent subgroups of similar character (for example, homologs). However, this would increase the already large computational burden of the EFDC code. It is unclear, however, the degree to which the contaminant transport portion of the model, which would need to be repeated for each contaminant subgroup, contributes to the computation as opposed to the hydrodynamic and sediment transport portions, which need be run only once.

Finally, it is unclear the degree to which the EFDC simulation of PCBs is even needed. The computation is to some extent duplicative but to another extent complementary with AQUATOX. In particular, the EFDC PCB simulation is proffered as a means to address the PCBs deposited in the distal floodplain. As discussed at length in my General Overview, the need to model explicitly the distal floodplain needs to be further evaluated through a coherent conceptual model. The alternative would be to model sediment fluxes alone with the hydrodynamic code, and use the predicted sediment fluxes as input to a separate PCB model (either AQUATOX or a simpler alternative).

5. What supporting data are required for the calibration/validation of the model on the spatial and temporal scales necessary to address the principal need for the model (as defined above)? What supporting data are required to achieve the necessary level of process resolution in the model?

The presumption of this question is that a comprehensive accounting of supporting data needs can be accomplished prior to data analysis and interpretation, conceptual model development, and even preliminary model development. In fact, data collection should go hand-in-hand with these other processes and be guided by them. For example, as discussed in my General Overview, there does not appear to have been an adequate attempt to analyze the data and understand the degree to which distal floodplains are a source of PCBs to the river system. The skeletal conceptual model outlined in my General Overview identifies a gap in understanding: we do not seem to know to what degree distal floodplains are a source of PCBs to the river and biotic system. Identification of this knowledge gap then leads to identifying such specific data needs as additional Sedflume studies specific to the distal floodplain and better characterization of flow dynamics during distal floodplain inundation.

With this general paradigm in mind, what needs to be done to answer Question 5 is to analyze the already available data with simple quantitative tools, develop preliminary conceptual models of important processes, and through those conceptual models, identify gaps in characterization and understanding. This process then leads logically to the identification of specific additional supporting data needs.

Although additional data analysis and conceptual model development is needed to identify data needs more completely, certain existing data needs can be identified. For HSPF, additional validation data sets would be valuable, but are probably not essential. The EPA may wish to consider a single tributary flow-gaging station with occasional TSS measures as supplementary data for HSPF validation. For EFDC, the lack of data to characterize a large storm event and floodplain inundation is an unfilled data gap. Although the occasion of this type of data is a vagary of nature, the EPA should allocate the resources to monitor such an event in detail if it occurs. As far as AQUATOX, as stated above, there are not enough data and probably never will be enough data to calibrate and validate a model of this complexity. Mr. Endicott identified selected data that would be valuable to collect. I concur with his recommendations, but in the final analysis, emphasis should be placed on developing and implementing a robust approach to uncertainty analysis for this component of the modeling framework.

6. Based upon your technical judgment, are the available data, together with the data proposed to be obtained by EPA, adequate for the development of a model that would meet the above referenced purposes? If not, what additional data should be obtained for these purposes?

The answer to Question 6 is much the same as the answer to Question 5: it is impossible at this stage in the study to judge whether the available data are adequate. What seems clear based on the peer review panel workshop and my own thinking about this question, is that the process being conducted by EPA appears to be inadequate. The specific inadequacy is premature selection of sophisticated and complex modeling codes without a clear prior understanding of the system dynamics and simple quantitative analysis to identify the importance processes within that system. I emphasize again the importance of developing a coherent conceptual model as discussed in my General Overview.

As far as specific data issues, I am concerned by the project team's failure to consider important data available from Massachusetts sources. The following potentially valuable data were not considered and apparently were unknown to the modeling team:

Data Source	Data type
Massachusetts Department of Environmental Protection	Past water-quality assessments in 1997-1998, 1992, 1985, 1976-1978, 1974, 1968-1969 that variously included water-quality sampling, wastewater discharge surveys, biota sampling, sediment sampling, and probably time-of-travel and other hydrodynamic field studies.
MassGIS	Geographic information system coverages of soil types, land use, wetlands, surficial geology, topography, aerial photography, and other geographical features.
Federal Emergency Management Agency	Hydrologic and hydraulic studies conducted under the Flood Insurance Program and possibly high-water mark surveys after flood events

These data are likely to assist in the formulation of food-web relations, construction and calibration of the hydrodynamic and hydrologic models, and construction and calibration of the phytoplankton component of the water-quality model. I am concerned that the failure to search for and incorporate these important past data could betray a false confidence in the proposed modeling framework: in other words, that the modelers could have concluded that their models are so good and so fundamentally sound, that they do not need to exert every effort to locate the best available data.

III. Specific Comments on the Modeling Framework Design Report and/or the Quality Assurance Project Plan.

- The QAAP does not appear to include procedures specifically to check model input data. Section 11 appears to touch on this, but should also specify that all input data time series be plotted for visual inspection and cross-checking.
- The QAPP does not include a discussion of the QA process governing code selection. The selection of particular codes should be appropriately documented by describing the alternative codes considered, the advantages and disadvantages of the alternatives, and finally how those factors were weighed in the final code selection. This should not be an after-the-fact apology for some predetermined codes: rather it should be an analytical and, if possible and appropriate, quantitative evaluation of the codes. Returning again to the example of the hydrodynamic model: it should be possible to estimate the relative magnitude of different transport mechanisms and thereby educate the selection of a one-, two-, or three-dimensional modeling code.
- The EPA Response to Peer Review Panelist Questions indicates that the new sediment bed representation in AQUATOX was tested against the IPX V 2.74 model. It, and all other newly created code, should be also validated against analytical solutions for which there are known solutions.

IV. Concluding Comments

In summary, I believe that the Modeling Framework Design for the Housatonic River is premature. Information presented and otherwise made available to the Peer Review Panel does not reveal that there is yet an adequate conceptual model of the Housatonic River system upon which to base the selection of modeling tools. While the HSPF model seems a logical and solid choice, the EFDC model is too burdensome computationally and the AQUATOX model too complex to ever be unequivocally calibrated. Misgivings over these choices are not assuaged by the information provided—a solid analysis of the data and river system does not appear to have been conducted to justify these choices.

Before the EPA fully commits to the modeling framework it has chosen, a coherent conceptual model is needed. This is particularly critical for the distal floodplains. It is not apparent based on the information given whether and to what extent the distal floodplains can act as a source of PCBs to the river channel and aquatic ecosystem. Until this fundamental question is answered, informed decisions as to the necessary dimensionality and spatial structure of the hydrodynamic model cannot be made.

The great complexity of the modeling framework necessarily implies considerable uncertainty in the calibration of the model parameters. This is particularly the case for the AQUATOX model. Although the complexity of AQUATOX seems excessive for this particular project, some level of complexity and uncertainty is unavoidable. Accordingly, it is imperative that the modeling framework be modified to include formal uncertainty analysis. To simply provide decision makers model predictions without information on the uncertainty of those predictions would be a disservice that could result in needless squandering of remediation funds. Decision makers need to be provided with a good assessment of the reliability and uncertainty of the model predictions so that choices between remediation alternatives are fully informed.

Cited References

Beach, R.B., P.M. Craig, R. DiNitto, A.S. Donigian, G. Lawrence, R.A. McGrath, R.A. Park, A. Stoddard, S.C. Svirsky, W.D. Tate, and C.M. Wallen, 2000a. Modeling Framework Design, Modeling Study of PCB Contamination in the Housatonic River. U.S. Army Corps of Engineers, New England District, Concord, Massachusetts and U.S. Environmental Protection Agency, Region I, Boston, Massachusetts. October 2000.

Beach, R.B., J.S. Clough, P.M. Craig, A.S. Donigian, R.A. McGrath, R.A. Park, A. Stoddard, S.C. Svirsky, and C.M. Wallen, 2000b. Quality Assurance Project Plan, Modeling Study of PCB Contamination in the Housatonic River. U.S. Army Corps of Engineers, New England District, Concord, Massachusetts and U.S. Environmental Protection Agency, Region I, Boston, Massachusetts. October 2000.

Beck, M.B., 1987. Water Quality Modeling: A Review of the Analysis of Uncertainty. *Water Resources Research*, Vol. 23, No. 8, Pp. 1393-1442. August 1987.

Beck, M.B., and E. Halfon, 1991. Uncertainty, Identifiability and the Propagation of Prediction Errors: A Case Study of Lake Ontario. *Journal of Forecasting*, Vol. 10, No. 1 & 2, Pp. 135-161.

Gailani, J.Z., S.J. Smith, M.G. Channell, G.E. Banks, D.B. Brister, R.A. Jepsen, and J.D. Roberts, 2000. Sediment Erosion Study for the Housatonic River, Massachusetts (Draft). U.S. Army Corps of Engineers, New England District, Concord, Massachusetts. September 2000.

Hartigan, J.P., J.A. Friedman, and E. Southerland, 1983. Post-audit of lake model used for NPS management. *Journal of Environmental Engineering, ASCE,* Vol. 109, No. 6, Pp. 1354-1370. December 1983.

Shanahan, P., and D.R.F. Harleman, 1984. Transport in Lake Water Quality Modeling. *Journal of Environmental Engineering, ASCE*. Vol. 110, No. 1, Pages 42-57. February 1984.

Tracy, J.C., and C.M. Keane, editors, 2000. Peer Review of Models Predicting the Fate and Export of PCBs in the Lower Fox River Below DePere Dam, A Report of the Lower Fox River Fate and Transport of PCBs Peer Review Panel. American Geological Institute, Alexandria, Virginia.

WDNR, 1999. Draft Studies Completed on Cleanup of PCBs in Lower Fox River Sediments. (Published newsletter.) Wisconsin Department of Natural Resources, Madison, Wisconsin. October 1999. http://www.dnr.state.wi.us/org/water/wm/lowerfox/rifs/foxrivercleanup.pdf, accessed April 18, 2001.

Weston, 2000. Supplemental Investigation Work Plan for the Lower Housatonic River. Roy F. Weston, Inc., Manchester, New Hampshire. Prepared for the U.S. Army Corps of Engineers, New England District, Concord, Massachusetts. February 2000.